Handwritten Gujarati Script Recognition

2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9441811

Archana Shirke
Department of Information Technology
Fr. Conceicao Rodrigues Institute of Technology
Vashi, India
archana.shirke@fcrit.ac.in

Nikunj Gaonkar
Department of Information Technology
Fr. Conceicao Rodrigues Institute of Technology
Vashi, India
Nikunj.gaonkar@gmail.com

Paresh Pandit
Department of Information Technology
Fr. Conceicao Rodrigues Institute of Technology
Vashi, India
pmpandit31086@gmail.com

Kapil Parab
Department of Information Technology
Fr. Conceicao Rodrigues Institute of Technology
Vashi, India
kapilaparab@gmail.com

***Abstract*—Their are numerous languages in this world and they are spoken and written every day. In modern times everyone has recognized the need for the recognition of handwritten scripts by machines for various purposes. Handwriting recognition has been one of the most demanding and interesting research areas in the field of image processing and machine learning since a long time. It finds its use in various applications which include conversion of handwritten documents into structural text form, reading aid for the blind, bank cheques, etc. Handwritten Script recognition can also play a significant role in preserving the culture, famous works, and literature from old times that were handwritten. For this reason, it is particularly important to have a handwritten script recognition system that can be used irrespective of the document type.**

The project aims at developing an offline handwritten Gujarati script recognition system which can be effectively used for recognizing handwritten Gujarati scripts. There is a significant difference in the handwriting of different people and the contours formed by characters of the Gujarati Script and these factors poses a stiff challenge in this process. This process involves the creation of a neural network that will take input as an image (.jpg,.jpeg, .png, etc.) and then the features will be extracted from it, the neural network will recognize this text and output it on the machine. This will help in the recognition of Gujarati scripts of different handwriting and will help reduce human error and save the time of individuals.

Keywords—Yolo; Text Recognition; Object detection; Handwriting Recognition; Gujarati Script Recognition;

I. INTRODUCTION

There are a large number of languages in the world and there is a lot of information/text present in the form of handwritten scripts which is generated every day during our day to day life. This caters to the need for automated recognition of handwritten scripts by a computer as there are a lot of risks associated with written material on paper such as tearing off or high susceptibility to weather conditions or maybe misplaced easily. Handwritten script recognition involves detection and recognition of the text and identifying specific words or sequence of characters from a piece of handwritten text. So developing such a system will help for automatic recognition of handwritten Gujarati scripts and will also eliminate the manual work involved which may contain errors, is time- consuming, and often tedious.

Today in the modern world when technology is evolving such rapidly, there is a high demand and need to convert manuscript or any kind of handwritten text in to a document in electronic/ machine readable format. It has become important to digitize documents. Also, as the trend of automation of official works is increasing it has promoted the researchers to bridge the gap between an average Person and the technology. Offline Handwritten Character Recognition, also known as HCR, is the process of recognition of handwritten, printed, or computer/typewriter printed text using computers. Character recognition systems are of two types: Offline and Online. Even today people are more often using pen and paper during their day-to-day activities, and so this has bought the attention of many researchers to offline character recognition There are numerous offline handwritten character recognition applications, such as blind reading aid, scene text recognition, electronic preservation of handwritten old/historical records, automated reading for postal mail sorting, bank cheques, automation of various administrative offices, etc.

Today in the data driven world where data is the king and more and more data generated every day, its importance is going to be ever-increasing, and the fact that over 90% of data in the world is generated just over last two years the importance of data can't be denied. Today there is a lot of useful information present in age old manuscripts and other handwritten documents but until and unless they are converted into Machine editable format it is extremely tough to analyze it and extract meaningful information from it. Therefore it's very important to have a system that can convert this huge handwritten data in machine editable format.

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

This offline handwritten Gujarati Script recognition system has a wide range of applications and scope. It will automatically recognize hand-written Gujarati Script accurately and efficiently, and it can be used to convert handwritten Gujarati text into a machine editable format/electronic format in quick time. It can also be used for various applications in sectors such as research, government, education, postal processing, scene text recognition, script recognition, etc.

The objectives of our system are as follows:

1. User Interface:

   A Desktop application that will allow users to upload an image and get the output txt file containing the recognized text.

2. YOLO model:

   The trained YOLO model is the backbone of proposed system that predict the output for input image and recognizesthe text.

II. LITERATURE SURVEY

A framework to recognize online handwritten character for the Gujarati dialect. The authors had majorly focused on using strokes to classify different classes. The authors have various process of character recognition such as input pre-processing, feature extraction in much detail. The overall dataset consisted of 3000 handwritten Gujarati images which were contributed by more than 100 writers. The strokes of each character are pre-processed before extracting features from them. For pre-processing methods such as normalization, smoothening and resampling are used. Feature extraction is done using zoning technique. By using SVM- RBF kernel accuracy of 91.63% was achieved and using MLP the accuracy was 86.72%.

Many researchers have been interested and contributed greatly in Handwritten Optical Character Recognition and this has been a field with constant improvement every year. Jayashree Prasad, Dr. U. Kulkarni, Rajesh Prasad (2009) have proposed a solution for gujarati recognition in the paper titled Offline Handwritten Character Recognition of Gujrati Script using Pattern Matching. They have combined template matching methods along with neural networks for recognition. This ensures the benefits of both the methods used in the implementation. They have achieved a 73.33% recognition rate for the 6 classes with a 71.66% efficiency rate. [2]

Prachi Solanki and Malay Bhatt (2013) have also presented on text based optical character recognition using classifiers. This follows the principle of segmentation for the sentence, word and character detection in the same order. The individual words are then sent to a trained Hopfield neural network classifier for the final output. They have achieved an accuracy of 93.2%. [3]

Gopal, Sanika, Sayali, Megha, Komal (2020), have proposed to use a Convolutional Neural Network for the digit recognition on handwritten numerical data. They have developed this system for the 10 digits from 0-9 in the gujarati akshar. They have been able to achieve an average accuracy of 95.4%. [4]

Handwritten Gujarati Script Recognition system based on Image Processing and Deep Learning. The main motive of the author of this paper is to highlight the key aspects of the proposed system which can detect and recognize handwritten Gujarati script by using machine learning and image processing. Detecting handwritten Gujarati characters is a difficult task due to vast difference in the handwriting of Individuals and the contours formed by the various characters of the Gujarati language. The authors have highlighted the complete process of character recognition. The project works by pre-processing the input image and then apply segmentation to first identify the lines and consequently the individual characters, finally deep neural networks are used for feature extraction to recognize the characters. It also explains in detail the process of training and testing of the model. They have also applied spell check for post processing the model output. [5]

'A review on different approaches used for Optical character recognition'. In this paper the authors have described two methods OCR using neural networks and OCR using correlation method. The correlation method works by pre-processing the input image first the image is converted to gray-scale and then it is converted to binary image subsequently noise filtration is applied. After pre-processing segmentation follows where individual characters are segmented from the complete image. Finally this image is compared with all the examples in the training set and the class with which the image shows the highest correlation score is considered as the class for that character. For performing OCR using neural networks the image is preprocessed as in the correlation method and then segmentation is performed to segment input images in to individual lines and further segmentation is done to extract individual characters from image. The neural network consists of two layers the training set consists of images with noise and without noise. [6]

'Gujarati Script Recognition: A Review' In this paper the authors have discussed the characteristics and various recognition techniques that can be used for recognizing various Indian scripts such as Devanagari, Gujarati and languages from South India, Gurumukhi script, Oriya and Bengali script and they have also explained in detail the various steps of handwritten Gujarati script recognition which includes data pre-processing techniques, segmentation of the input image, techniques for extracting features from the image and lastly how recognition is done by the classifier. [7]

III. PROPOSED SYSTEM

The proposed system will be used to detect and recognize Gujarati handwriting on scanned images. The first step is to create a dataset of images of handwriting samples that are required for the training of the model. The data collection process consists of making a dataset of handwritten Gujarati images from various sources available, labeling, noise

removal, and outlier removal. The approach at labeling will be to first make an appropriate dataset of handwritten Gujarati images and then the dataset would be labeled using our customized LabelImg tool which supports many advance functionalities to serve the purpose of labeling. We aim at designing a tool with which a user can generate his own training data and train the model themselves and test it or use our recognition model.

Once the dataset has been labeled correctly, a neural network will be developed for recognition of the handwriting. We have considered using YOLO v3 as it has increased accuracy rate over the other versions. YOLO algorithm will be used for detection. YOLO stands for You Only Look Once and follows a single-shot object detection method. It is comparatively fast in object detection, and it uses Darknet-53 for feature extraction. The Neural Network will be trained on the scanned dataset usingfeature extraction enabled by default. The task of extracting features from the image is done by the initial convolutional layers of the network while the fully connected layers predict output probabilities and coordinates. Once the cost function shows minimal loss the training can be stopped.

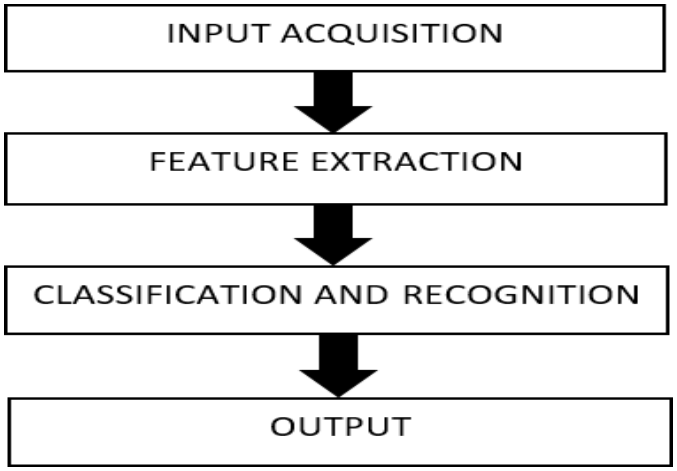

Fig.1. Flow diagram of proposed system

The system mainly consists of the following three modules:

A. Data Collection:

Here we collect the images of Handwritten Gujarati Images from various sources and annotate the *swar*, *akshar* and *maatras* using our customized LabelImg tool that is specially designed to meet the purpose of our system. Users can also generate their own training data so that they can use it for training if they want.

B. Training:

The Gujarati text images that are annotated in YOLO format could be then fed to our training module where we have used a YOLO algorithm to extract features from the images and learn the generalized representations of the object and to encode the contextual information from the image so as to make the system better and better as the user trains it with more data of his own making our model more robust to detect different handwriting.

C. Testing:

Using our testing module, a user can test our model for its accuracy using a pre-trained weights file, the user inputs a single image at a time so that the model/system starts to predict the characters and therecognized text is stored in a txt file at a desired location chosen by the user.

IV. Implementation Details:

A. Labeling Tool:

YOLO is a single shot image detection algorithm, in order to train a model using YOLO we need to annotate data in YOLO format, an image annotated in YOLO formatgenerates a txt file that consists of class_index, x_cen (x co- ordinate of the center of the bounding box), y_cen (y co- ordinate of the center of the bounding box), width and heightof the bounding box. We have used a labeling scheme for all the *swar*, *akshar*, and *sanyukt akshar* according to the scheme and their respective UTF-16 encoding.

Since there are a lot of literals in Gujarati Script resulting from a combination of *swar*, *maatra* we have around 820 classes in our customized LabelImg tool.

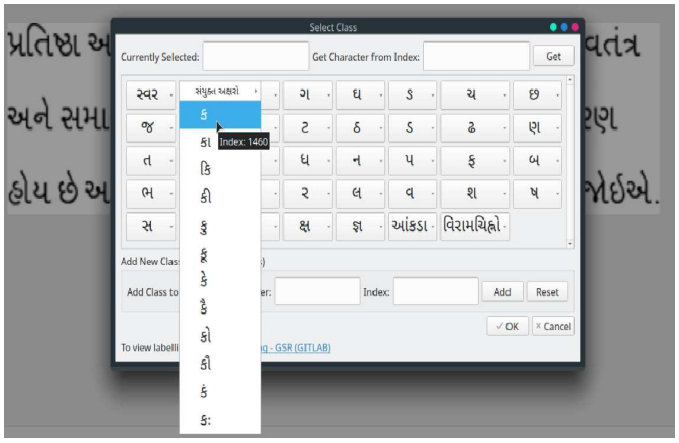

Fig. 2. The labeling tool

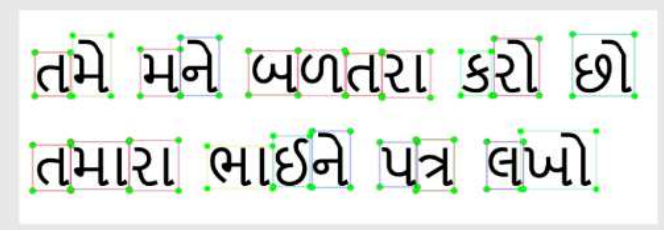

Fig. 3. Annotated characters

B. Generating the dataset

For the generation of training dataset, we have collected data from various online sources that include newspaper clippings, web scraping of openly available images of articles, magazines and posters and legal document. The annotations text file that is generated contains class_index, yolo coordinates, and the file names. This also contains metadata about the image foradditional data inference such as the path, and number of bounding boxes in the image that are present in the files but will not be used for the training. The class consist of the 26 alphabets in capital and small along with 0-9 digits and standardpunctuation signs.

In order to achieve a proper standard of detection, we need large amounts of labelled data. Thus apart from annotating data manually using our tool which is a bit time consuming we also tried to automate the process of annotating the dataset. For this, we have used a library named trgd available as a pypi package. This takes input a dictionary of all the characters that we need to use for generating single line sentence images with adequatewords and enough iterations to cover all occurrences of the input characters. This can eliminate any imbalance in the training dataset caused by less data for a particular class. Using the input labels, and contour detections, we can get the annotations as per the YOLO requirements. We have selected a kernel size for the method used to detect the contours which will help in theauto labeling process.

We have collected many images of different sizes, multiple fonts and orientation along with multi line and single line images to generate enough annotations such that we can get a good detection after training the YOLO model. As such we have annotated a more than thousands of images to generate enoughdata to train our model.

```
20 0.018914 0.506757 0.027961 0.527027
10 0.043174 0.574324 0.020559 0.716216
39 0.063322 0.520270 0.013158 0.527027
32 0.087582 0.506757 0.030428 0.527027
29 0.132400 0.581088 0.023026 0.405405
```

Fig. 5. Character labelled annotation text

C. Characteristics of the proposed Deep Learning Model:

YOLO (You Only Look Once) is considered by many as one of the best object detection algorithms as compared to its counterparts. The algorithm "only looks once" at the image means that the algorithm requires just a single forward propagation pass of the image for the network to make predictions. YOLO sees the complete image at the time of training and testing so it implicitly extracts complete information about classes as well as their appearance and learns the features of classes. It learns general representations of objects so that when it is trained extensively and tested on the other images, the algorithm outperforms other top detection methods. YOLO is a fully convolutional network (FCN) as it is consists of all convolutional layers. YOLO uses Darknet-53 as its future extractor, as its name suggests, it contains 53 convolutional layers, each followed by a batch normalization layer and Leaky ReLU activation. Pooling is not used here instead a convolutional layer with a step value of 2 is used to down sample the feature maps. This helps to preventing the loss of low-level features often that is caused by pooling. It is a fully convolutional network and its final output is generated by applying a 1 x 1 kernel on a feature map, the detection is done by applying 1 x 1 detection kernel on feature maps of three varied sizes at three separate places in the network. The 52 x 52 layer detects small size objects, and the 26 x 26 layer detects medium size objects and the 13 x 13 layer detects the small size objects.

For making predictions the input image is divided into S X S grids, the grid cell which appears on top of grid cell, is responsible for the detection of that object. All such grid cells predicts multiple bounding boxes and a confidence score for each of the bounding box, this confidence score indicates that how much correct is the bounding box that the YOLO has predicted and then max pooling is applied to remove multiple bounding box for same object that eliminates bounding box with a confidence less than certain threshold value. Each bounding box has 5 values associated to it x_cen, y_cen, w, h and confidence score, where x and y are the coordinates of the center of the bouding box, w and h refers to the width and height of the detected bounding box respectively.

During the training we have used 450 images with around 100- 150 characters in each image and the overall dataset contains around 500 samples of each class the model was trained for a duration of 120 hours on a NVIDIA GPU with 8 GB VRAM.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Fig. 5. Network Architecture

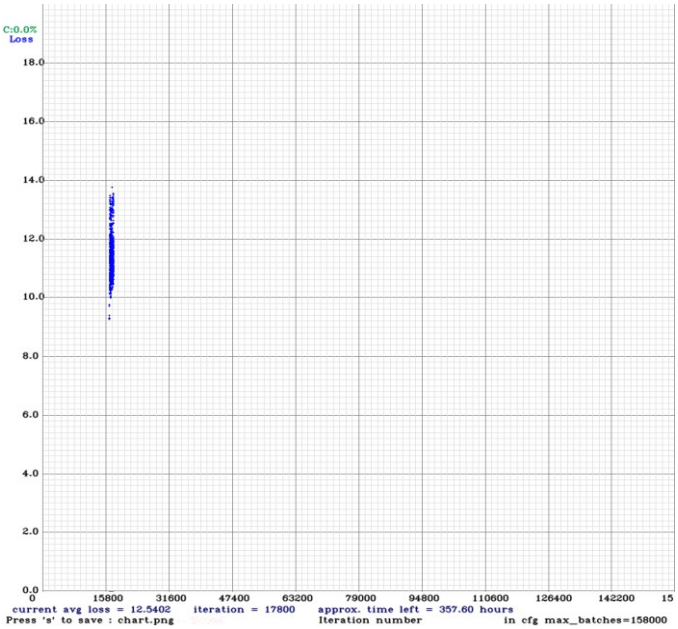

Fig. 6. Loss chart for training

D. Deployment:

The end product will be a desktop client that is written in python using the PyQT library which is a wrapper for the QT GUI library and it allows for robust GUI development including cross platform support. The standard features include uploading of the image for the conversion, and the text output. Along with the standard features it will also allow to load custom trained models for testing purposes. The appimage is built using the AppImageKit which requires OpenCV4 for the execution. The appimage is portable and does not require any installation procedure to run.

V. Result Analysis

The proposed is aimed at detecting and recognizing handwritten Gujarati characters, text/script by using the YOLO darknet framework and various deep learning techniques, and results are returned in a machine editable format i.e. a txt file. Recognizing each and every character of the script accurately at an improved speed while reducing the time required for conversion, has been a major goal of the proposed system. The software generates an output text file which is stored in the users machine at a location choosen by them. The results we achieved shows that the YOLO algorithm also works well for recognizing handwritten text. In order to form a proof of concept for working with Gujarati script we have first developed the model for English Language and we will use the same procedure for Gujarati recognition.

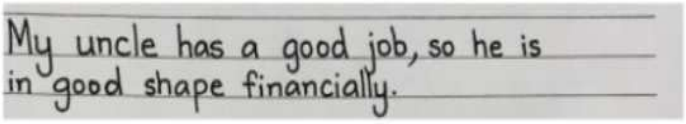

Fig. 7. Input image

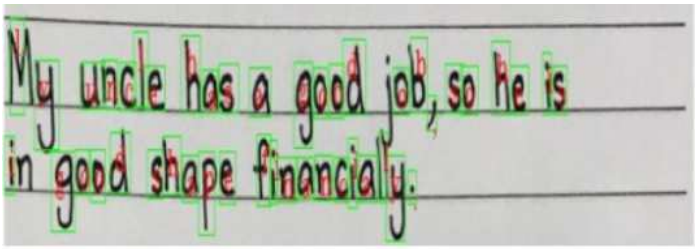

Fig. 8. YOLO detections on input image

VI. Conclusion

The task of recognizing handwritten Gujarati script is quite challenging as there is a high structural complexity in the characters of Gujarati language, the variation due to different styles and handwriting of individuals and the poor quality of source image/document degradation over time. Moreover the contours formed by the characters of the Gujarati language also have quite a lot of variation which makes detection a complex process and sometimes a minor change or extension of stroke that causes a difference between two letters. This system has an advantage that the accuracy is improved, speed is increased and time is reduced as we are using the YOLO darknet framework, on the further training of the neural network with a huge training dataset, the various characters, digits and punctuations of the Gujarati Script recognition results would be further improved to a great extent. The

advantage of using this system is that it uses the YOLO algorithm which makes it faster while not compromising accuracy at all, and the fact that the project has add-on functionalities that users can make use of so as to achieve even better results by training the model on custom dataset. The success of this proposed system is a significant aid to this field of research. The proposed system can be used to convert Gujarati manuscripts as well as the handwritten Gujarati documents to a machine readable format.

Acknowledgment

We would like to convey our admiration to everyone who provided us the resources and motivation to complete the project. We respect and thank Dr. S. M. Khot, Principal of Fr. C Rodrigues Institute of Technology, Prof. Dhanashree Hadsul, HOD/Co-ordinator of the I.T. Department for extending their invaluable support to us. We are also grateful to our guide Mrs. Archana Shirke who has supported us throughout our project period and also provided us with all the important information required for developing this project. We consider ourselves fortunate enough to be encouraged and guided by all the teaching as well as non-teaching staff of the I.T. Department. Also, we would like to extend our warm thanks to our family members and well-wishers for always supporting and encouraging our endeavor.

References

[1] Apurva A. Desai, Vishal A Naik, "An online handwritten Gujarati character recognition system based on SVM, MLP, KNN algorithm', 2017 International Conference on Communication, Computing and Networking Technologies (ICCCNT)

[2] Jayashree R. Prasad, U. V. Kulkarni, Rajesh S. Prasad, "Offline Handwritten Character Recognition system based on Pattern Matching Algorithms", 2009 International Conference on Anti-counterfitting, Security, and Identification in Communication

[3] Malay Bhatt and Prachi Solanki, "Typewritten Gujarati Script OCR using Hopfield Neural Network", International Journal of Computer Application

[4] G.D. Upadhye, Sanika M Shirsat, Sayali R Shinde, Megha A. Sonawane, Komal D. Pandit, "Pattern Recognition using Convolutional Neural Network for Handwritten Gujarati Numerals", International Journal of Recent Technology and Engineering (IJRTE)

[5] S. Aniket, R. Atharva, C. Prabha, D. Rupali, P. Shubham, "Handwritten Gujarati Script Recognition System with Image Processing and Deep Learning", 2019, International Conference on Nascent Technologiies in Engineering. (ICNTE)

[6] Pranob K Charles, Ch. Deepthi, V. Harish, M. Swathi, "A review on various techniques used for the Optical Character Recognition", International Journal of Engineering Research and Applications.

[7] Mamta Maloo, Dr. K.V. Kale, "Gujarati Script Recognition: A Review", 2011 International Journal of Computer Science Issues, Vol. 8, Issue 4 (IJCSI)